

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-134188

(43)公開日 平成9年(1997)5月20日

(51)Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
G 1 0 K 15/04	3 0 2		G 1 0 K 15/04	3 0 2 D
G 1 0 H 1/00			G 1 0 H 1/00	Z
	1 0 2			1 0 2 Z
G 1 0 L 3/02			G 1 0 L 3/02	A

審査請求 未請求 請求項の数4 O L (全11頁)

(21)出願番号 特願平7-292403

(22)出願日 平成7年(1995)11月10日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)発明者 林 克彦

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 森 大輔

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

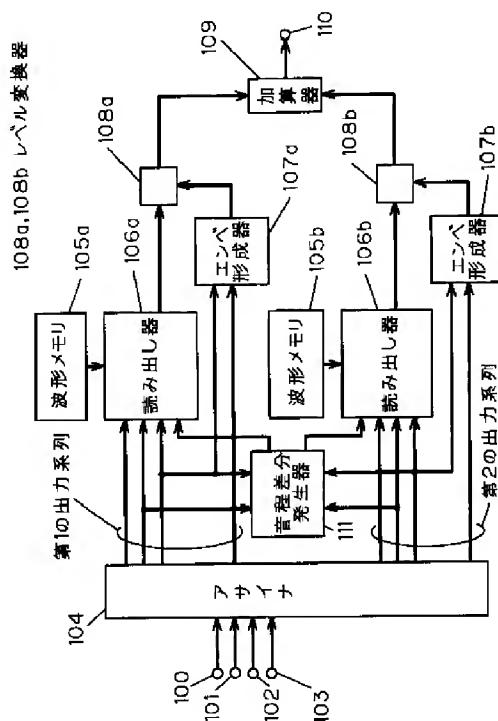
(74)代理人 弁理士 滝本 智之 (外1名)

(54)【発明の名称】 歌声合成装置および音楽再生装置

(57)【要約】

【課題】 キーボンによって音程を変化してもバックコーラスの品質が低下しない歌声合成装置および音楽再生装置を提供する。

【解決手段】 音韻と音高と音量および発音タイミングを第1、第2の読み出し器と第1、第2のエンベ形成器に順に与えるアサイナと、アサイナからの指示で波形メモリからデータを読み出す第1、第2の読み出し器と、アサイナからの指示で波形包絡データを生成する第1、第2のエンベ形成器と、第1、第2の読み出し器の出力を第1、第2のエンベ形成器の出力でレベル変換する第1、第2のレベル変換器と、第1、第2のレベル変換器の出力を加算する加算器を備えた歌声合成器であり、指定された音韻／音高の音声波形を、音韻毎に個別に読み出し器で読み出し、クロスフェードしながら滑らかに出力するクロスフェード部以外は原音PCMそのままで品質の高い歌声が合成できる。



【特許請求の範囲】

【請求項1】 入力された音韻と音高と音量および発音タイミングからなる演奏情報を第1および第2の出力系列に交互に出力するアサイナと、前記アサイナの第1および第2の出力系列によってそれぞれ音高と音韻と発音タイミングを指示される第1および第2の読み出し器と、前記アサイナの第1および第2の出力系列によってそれぞれ音量と発音タイミングを指示され波形包絡データを生成する第1および第2のエンベ形成器と、予め複数の音韻および音高の音声波形を記憶していくそれぞれ第1および第2の前記読み出し器によってデータが読み出される第1および第2の波形メモリと、前記第1および第2の読み出し器の出力をそれぞれ前記第1および第2のエンベ形成器の出力でレベル変換する第1および第2のレベル変換器と、前記第1および第2のレベル変換器の出力を加算する加算器とを備えた歌声合成装置。

【請求項2】 入力された音韻と音高と音量および発音タイミングからなる演奏情報を第1および第2の出力系列に交互に出力するアサイナと、前記アサイナの第1および第2の出力系列によって音高と発音タイミングを指示されて第1および第2の出力にそれぞれ第1および第2の音程変化値を出力する音程差分発生器と、前記アサイナの第1および第2の出力系列によってそれぞれ音高と音韻と発音タイミングを指示され、かつ前記音程差分発生器の第1および第2の出力によってそれぞれ音程変化値を指示される第1および第2の読み出し器と、前記アサイナの第1および第2の出力系列によってそれぞれ音量と発音タイミングを指示され波形包絡データを生成する第1および第2のエンベ形成器と、予め複数の音韻および音高の音声波形を記憶していくそれぞれ第1および第2の前記読み出し器によってデータが読み出される第1および第2の波形メモリと、前記第1および第2の読み出し器の出力をそれぞれ前記第1および第2のエンベ形成器の出力でレベル変換する第1および第2のレベル変換器と、前記第1および第2のレベル変換器の出力を加算する加算器とを備えた歌声合成装置。

【請求項3】 楽器および歌唱の演奏情報を記憶している演奏データメモリと、前記演奏データメモリから読みだした演奏情報と移調入力および演奏速度入力に応じてタイミングを制御しながら演奏指示情報を出力するシーケンサと、前記シーケンサの演奏指示に従って楽音を合成する楽音合成装置と、前記シーケンサの演奏指示に従って歌声を合成する歌声合成装置と、前記楽音合成装置の出力と前記歌声合成装置の出力を加算する加算器を備えた音楽再生装置。

【請求項4】 歌声合成装置は、請求項1または2に記載の構成であることを特徴とする請求項3に記載の音楽再生装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はバックコーラス付きのカラオケ演奏を行うカラオケ装置に用いる音楽再生装置に関し、特に曲のテンポ(演奏速度)およびピッチ(移調)変更によってバックバックコーラスの品質を低下することのない、歌声合成装置および音楽再生装置に関する。

【0002】

【従来の技術】近年、デジタル技術の進歩に伴い、各種のカラオケ方式が提案されており、特に電子楽器と同様の楽音合成装置を用いたシンセサイザ方式のカラオケ(以下、シンセカラオケと称す)が注目されるようになってきている。

【0003】シンセカラオケは、楽音合成装置を一般にMIDIと呼ばれる演奏指示情報等によって駆動し、カラオケ用のバックミュージックを発生するものであり、一般的のカラオケ(例えばレーザーディスクのカラオケ)に較べて、1曲の演奏に要するデータ量が非常に少ないと特長がある。

【0004】シンセサイザ方式のこの特長は、演奏情報をISDNや電話回線を伝送路として転送するシンセカラオケ(いわゆる、通信カラオケ)や演奏情報を半導体メモリ等に格納するタイプのシンセカラオケ等に生かされていて、今後この分野でのシンセカラオケの成長が期待されている。

【0005】反面、シンセサイザ方式のカラオケには音声やバックコーラスの演奏が難しいという欠点が存在する。これは、シンセカラオケが電子楽器の楽音合成装置を用いて構成されており、楽音合成装置が楽器音(ピアノやギター等の音声以外の楽器)の再生を目的として構成されていること起因している。

【0006】このような、欠点を克服するために従来考案されたシンセサイザ方式のカラオケの音楽再生装置として、例えば、特開平6-161479号公報に開示されたものがある。

【0007】この従来の音楽再生装置を同公報記載の図面を用い、以下、その説明を行う。同公報の図1において1は楽音データ記録部、2は楽音演奏部、3は入力部、4は文字データ記録部、5は音声合成部、6は合成増幅部、7は出力部、8は同期部である。

【0008】以上のように構成された従来の音楽再生装置について以下説明を行う。楽音データ記録部1は演奏曲の録音データを記録するもので、例えばコンパクトディスク(CD)、レーザーディスク(LD)または磁気記録媒体とそのドライブ装置からなる。楽音演奏部2は楽音データ記録部1の楽音データを取り、アナログの楽音信号に変えるものである。ここで記録部1に記録されている楽音データをMIDIコードとすれば、楽音演奏部2は電子発振器を使って音を生成するシンセサイザを用いればよい。また、入力部3は歌唱者の音声をマイク等で入力する。また、文字データ記録部4は文字コード列

を記録するもので、楽音データ記録部1と同様に磁気記録媒体等とドライブ装置から構成され、楽音データ記録部1と兼ねることも可能である。音声合成部5は文字データ記録部4に記録された文字データ列を受取音声信号に変換するものである。また合成增幅部6は楽音演奏部2の出力と入力部3より入力された音声信号と音声合成部5の出力を合成し、かつ増幅する。その結果の出力はスピーカ等の出力部7に送られる。また同期部8は楽音演奏部2の演奏に合わせて文字データ記録部4から音声合成部5へ送られる文字データの制御を指令する。

【0009】音楽データ記録部1にはMIDIコードが記録されているものとする。同公報の図2は文字データ記録部4に記録されている文字データの例を示す。図

1、図2を参照して本装置の動作を説明する。MIDIコードが演奏され始めると、楽音演奏部2に設定されている楽音データ記録部1からの読みとり速度、すなわち楽音演奏速度の値が同期部8へ送られ、同期部メモリ9へ記憶される。次にMIDIコードの第1小節のデータは楽音データ記録部1から読出され、楽音演奏部2に送られて楽音信号に変換される。同時に1小節分のMIDIコードを受け取ったことは同期部8へ知らされる。この同期部8は1小節のMIDIコードが楽音演奏部2に送られたことを知ると、文字データ記録部4から1行データを読出して文字数をカウントし、音声の速さ指定の値を計算する。ここでは1小節分の文字データは改行までの1行としている。そして音声合成部5は速さの指定から発音する文字コードデータを送る。速さの値は次のように演算される。速さの値 = (1小節の時間) / (1行の文字数)。

【0010】例えば、いま演奏しようとしている曲が1小節5秒の速さで演奏されるように設定されたとし、バックコーラスの1行の文字数が10文字であったとすると、音声合成部5は指定された速さで送られた文字を音声信号に変換し、合成增幅部6へ送る。この結果、楽音演奏部2で発生された楽音と、音声合成部5で発生された音声との同期がとられ、合成增幅部6より出力部7へ送られる。このように音声データを文字コード列として記録し、その文字コード列から音声を発生させるようにしているので、音声データも楽音のMIDIデータと同様にコード化して記録することができ、大幅にデータを圧縮することができる。

【0011】しかしながら、ここに開示された音楽再生装置では、文字データ記録部4には図2に示したように文字のみが記憶されていて、これらの文字を音韻に変換することを音声合成部6に指示することはできるが、音声合成部6に対して音程を指示することができないうえ、これらの音韻の発音タイミングも1小節内で均等に割り付けられるため細かく制御することができないなどの制約があり、特殊な楽曲用のバックコーラスしか実現できないという課題があった。

【0012】この点を改良した従来のシンセカラオケの音楽再生装置として、例えば、特開平7-140991号公報に開示されたカラオケ装置がある。

【0013】以下、従来の音楽再生装置を同公報の図面を用いて、以下に、その説明を行なう。

【0014】同公報の図1において、1はCPU、2はRAM、3はバス、4は曲データ記憶装置、5はパネルI/F、6はコントローラ、7は背景映像記憶/再生部、8は画像/歌詞表示部、9はビデオセレクタ、10はモニタ、11はマイク、12はミキサおよびエフェクタ、13はアンプおよびスピーカー、14は音源、15はプログラムROMおよびシーケンサ、16はディジタル音声デコーダである。

【0015】同公報の図2は曲データの構造図であり、図3は演奏データトラックの詳細を示す説明図であり、図4は音声データ指示トラックの詳細を示す説明図であり、図5は各トラックの時系列的な構造図であり、図6は音声データをデコードする部分の説明図であり、図7はコーラス音声制御のタイムチャートである。

【0016】以上のように構成された、従来のカラオケ装置について、以下その動作を説明する。

【0017】図1は従来例のカラオケ装置全体のブロック図である。この図において1はシステム全体の動作を制御および管理するCPU(中央処理装置)、2はCPU1によりシステム全体の動作を制御および管理するときに使用されるRAM(ランダムアクセスメモリ)、3はシステム全体を統合するためのデータおよびアドレスバスである。

【0018】4は複数の曲データを格納した記憶装置、30例えばHDD(ハードディスク)、5はパネルI/F(インターフェース)、6はパネルI/F 5を介してシステムに指示を与える複数のリモコン等のコントローラ、7は背景映像記憶/再生部、8は背景の静止画や歌詞を表示するための画像/歌詞表示部、9は背景映像記憶/再生部7からの背景映像(動画)と画像/歌詞表示部8からの画像を選択、合成するビデオセレクタ、10はビデオセレクタで選択、合成された画像を表示するモニタである。

【0019】11は歌い手の歌唱音声を入力するマイク、12は歌唱音声と演奏曲の楽音とを混合するミキサ40および各種の音響効果を与えるエフェクタである。13は合成された歌唱音声と演奏曲をアンプで増幅して出力するスピーカである。14は演奏曲の楽音を発音する音源、15はこの音源14並びに、ミキサおよびエフェクタ12を制御するシーケンサである。このシーケンサ15はCPU1が使用するプログラムを格納したプログラムROMを有する。16は符号化されたディジタル音声データ(例えばPCMやADPCMデータ)をデコードする音声デコーダである。

【0020】以下、動作を説明する。コントローラ6の操作によって演奏曲が指定されるとCPU1は記憶装置

4に記憶されている曲目リストを参照して該当する演奏曲の曲データおよびバックコーラス用の符号化された音声データをRAM2に転送し、シーケンサ15に制御を移す。シーケンサ515は、曲データに含まれる複数のイベントデータを基に、曲演奏を含む複数のイベントを同時に並列して実行する。

【0021】即ち、シーケンサ15は、楽曲の音色および疑似音声の音色に関するデータは音源14に、また符号化された音声データはディジタル音声デコード16に、さらに背景映像番号は背景映像記録/再生部7に、また歌詞番号は画像/歌詞表示部8にそれぞれ供給する。この結果、モニタ10の全面に背景映像が表示され、その一部に歌詞がスーパーインボーズされた表示状態となる一方、スピーカ13からは演奏曲およびバックコーラスが出力される演奏状態になる。

【0022】曲データは図2に示すように、ヘッダ部、データシーケンス部、音声データ部からなる。ヘッダ部にはその曲固有の情報として、曲番号、曲名、作曲者名、歌手名、背景画像選択情報、歌詞のフォント情報が書き込まれている。

【0023】データシーケンス部には、並行して同時に実行される複数種類のイベントを記述した複数のトラックが設定されている。音声データ部にはデータシーケンス部の音声データ指示データで選択される複数の音声番号が書き込まれている。

【0024】データシーケンス部に含まれるトラックの内、演奏データトラックには、音源14から演奏曲の楽音を発声させるデータを時系列に記述してある。疑似音声データトラックには、音源14から疑似コーラス音声(例えば「わー」、「うー」)を発声させるデータを時系列に記述してある。音声データ指示トラックにはデコード16でデコードされる真性コーラス音声(例えば「はこだてー」、「ながさきー」)の種類を指示する音声データ番号、音程、音高の各データが記述してある。歌詞データトラックには、演奏に併せてモニタ10に表示する歌詞の種類を指示するデータを時系列に記述してある。効果制御データトラックには、ミキサおよびエフェクタ12を制御するコントロールデータを時系列に記述してある。

【0025】図3は演奏データトラックの詳細を示す説明図である。この演奏データトラックには、ノートイベント、音色変更イベント、ピッチペンドイベントの各情報が記述されている。ノートイベントには、音源14の発声させようとする一つのCH(チャンネル)を指示するCHナンバ、ノート番号(音高)、ベロシティ(音量)、符長が書き込まれている。音色変更イベントにはCHナンバ、音色データが書き込まれている。ピッチペンドイベントには、CHナンバ、ピッチペンド情報が書き込まれている。演奏データトラックのノートイベントは、演奏曲の楽音発音用であるが、疑似音声データトラックも、

CHナンバが異なるだけの同様構造のノートイベントを有する。

【0026】図4は、音声データ指示トラックの詳細を示す説明図である。この音声データ指示トラックには、音声指示イベントの各情報が記述されている。即ち、音声データ番号、音程、音量の各情報である。音声データ番号は、デコード16でデコードされる符号化された真性コーラス音声データの番号であり、図2の音声データの音声番号のことである。予め複数のトラックは図5に示すように、イベントの種類と次のイベント発声までの待ち時間Δtを時系列に配列した構造である。

【0027】図6は音声指示イベントの情報に従い、音声データをデコードする部分の説明図である。シーケンサ15は、音声指示イベントの音声データ番号を用いてRAM2から該当する音声データ番号の符号化されたディジタル音声データを読み出し、それをディジタル音声データは、従来例では真性コーラス音声データである。一例としてディジタル音声データが、データ容量を減らすために圧縮されたアダプティブ・デルタ(ADPCM)データである場合、デコード16はビット数変換および周波数変換を行って伸長する機能を有したPCMデコードである。

【0028】デコード16の後段には音程および音量を制御するプロセッサ17が配置され、ここでデコードされたアナログ波形の音程および音量を、音声指示イベントに含まれる音程および音量情報に基づいて調整した後、ミキサおよびエフェクタ12に入力する。この経路を用いて真性コーラス音声を再生するのは再生速度が100%のときまたはピッチの変更量が所定範囲内のときである。コントローラ6からの指示で演奏曲の再生速度を100%以外に変更したときはプロセッサ17においてデコード16の出力に対する音量を0にして、真性コーラス音声の再生を中止する。ピッチの変更量が所定量の200~300セトを超えた場合も同様である。

【0029】ピッチ変更時に真性コーラス音声を使用しないのは次の理由による。つまり、音程、音量制御部17において、真性コーラス音声に対する音程をある程度以上変化させると、音声の質が低下し、実質的に使用できなくなる。「ある程度」の目安は半音単位で±2~3音(200から300セト)である。したがって、テンポ(再生速度)を変更した場合だけでなく、ピッチ(音程)を所定量以上変更したときも疑似コーラス音声を使用する。

【0030】以上の理由から、演奏曲の再生速度が100%以外になったとき、またはピッチの変更量が所定量の200~300セトを超えたときは、シーケンサ515は疑似音声データトラックのデータに従い、音源514から疑似コーラスを発声させる。疑似コーラス音声は、予め音源514に、楽音と同様にして「わー」「う

一」という音声のチャンネルを設定しておき、それを疑似音声データで選択して発音させるものである。

【0031】図7は、テンポ(再生速度)が100%を中心変化した場合の、シーケンサ15による制御の様子を示している。この例では音声データトラックのイベントは、時刻t1、t5、t8で順番に変化する。これに対しテンポは時間t4で100%から70%に、時刻t6で70%から120%に、時刻t7で120%から100%に変化した例である。

【0032】時刻t1のテンポは100%なので、この時は音声データトラックのイベントに従い、「はこだてー」という真性音声データのデコードを開始し、音声データの音量を音声イベントに記述された所定量に設定する。このとき、疑似音声データトラックについては何も処理しないという訳ではなく、最初の「は」に相当する音声イベントを読み、その疑似音声(例えば「わー」)を音源14から発音させる。但し実際には、疑似音声側の音量を0にして、スピーカ13からは出力できないようにする。同時に、真性音声データに関する音量値(上記の所定量)を復帰用に記憶しておく。

【0033】時刻t2でもテンポは100%なので、真性コーラス音声の出力を継続する。このとき、疑似音声トラックは「こ」の発音イベントに入るが、疑似音声側の音量は0のままである。時刻t3でもテンポは100%なので、真性コーラスの音声の出力を継続する。このとき、疑似音声トラックは「だ」の発音イベントに入るが、疑似音声側の音量は0のままである。

【0034】時刻t4になるとテンポが100%から70%に下がるので、真性音声をフェードアウトし、代わりに疑似音声の発音チャンネルを予め記憶しておいた所定量にフェードインする。このときは、真性音声が再生途中である可能性があるので緩やかなクロスフェードを行って途切れを回避する。

【0035】時刻t5になると音声データ側は次のイベント(「しんじることさ」)に入るが、テンポが100%に戻っていないので、音量を0にしたまま真性音声の発音はせず代わりに「し」に対する疑似音声(例えば「わー」)を音源14から発音させる。この疑似音声については、イベントに記述されたCH番号、ノート番号(音高)、ペロシティ(音量)、符長を用いて制御する。

【0036】時刻t6でテンポが70%から120%に変化するが、100%ではないので「ん」に相当する疑似音声(例えば「わー」)の発音を継続する。時刻t7でテンポが70%から100%に変化するので真性音声に戻る1つの条件が成立したが、この段階では前の音声データイベントのデコード途中なので、次の音声データイベントが開始するまでは、疑似音声の発音を継続する。

【0037】時刻t8になるとテンポが100%であり、かつ新たな音声データイベントが開始するので、このタイミングで真性音声にフェードインさせる。この時

は真性音声の開始時なので、急峻にフェードインしても問題ない。この時、疑似音声の「ど」に相当する発音部分もあるが、その発音チャンネルの音量を0にして発音を禁止する。以上のような制御をシーケンサ15が行うことことで、コーラスパート付きのカラオケ演奏曲を再生速度が変化しても常に時間のずれの無い状態で提供することができる。

【0038】

【発明が解決しようとする課題】しかしながら上記の従来の構成では、演奏速度を変化したり音程を変化するとPCMおよびPCM圧縮波形で形成されたバックコーラスから楽音合成装置で形成された疑似バックコーラスに切り替わるため、音質が変化してしまい、特に演奏速度と音程が規定値内ではないときは、疑似バックコーラスになって、バックコーラスの品質が低下するという誤題を有していた。特にPCMおよびPCM圧縮で形成したバックコーラスが歌詞を再生するのに対し、「あー」、「わー」で構成した疑似バックコーラスは異質である。

【0039】また、PCMおよびPCM圧縮形式のバックコーラスのデータ量は、楽音合成装置を駆動する一般にMIDIと呼ばれる演奏情報を較べて、非常に多い(100倍以上)ため、通信カラオケ等ではこれらを曲毎に通信回線を経由して伝送するのに多大の時間を要し、またスタンダードアロンのシンセカラオケ等では大量のデータ記憶装置が必要になる。これを避けるためには演奏速度または音程が規定値内であっても疑似バックコーラスを用いるようにしなければならないという課題を有していた。

【0040】本発明は上記従来の課題を解決するもので、PCM品質のバックコーラスを合成できる歌声合成装置を提供し、さらにこれをシンセカラオケに適用することにより、演奏速度や音程が変化してもバックコーラスの品質が低下しない音楽再生装置を提供するものである。

【0041】

【課題を解決するための手段】これらの課題を解決するために、本発明の歌声合成装置は、入力された音韻と音高と音量および発音タイミングからなる演奏情報を第1および第2の出力系列に交互に出力するアサイナと、アサイナの第1および第2の出力系列によって音高と発音タイミングを指示されて第1および第2の出力にそれぞれ第1および第2の音程変化値を出力する音程差分発生器と、アサイナの第1および第2の出力系列によってそれぞれ音高と音韻と発音タイミングを指示されかつ音程差分発生器の第1および第2の出力によってそれぞれ音程変化値を指示される第1および第2の読み出し器と、アサイナの第1および第2の出力系列によってそれぞれ音量と発音タイミングを指示され波形包絡データを生成する第1および第2のエンベ形成器と、予め複数の音韻および音高の音声波形を記憶していてそれぞれ第1および

第2の読み出し器によってデータが読み出される第1および第2の波形メモリと、第1および第2の読み出し器の出力をそれぞれ第1および第2のエンベ形成器の出力でレベル変換する第1および第2のレベル変換器と、第1および第2のレベル変換器の出力を加算する加算器とを備える。

【0042】また、本発明の音楽再生装置は、楽器および歌唱の演奏情報を記憶している演奏データメモリと、演奏データメモリから読みだした演奏情報と移調入力および演奏速度入力に応じてタイミングを制御しながら演奏指示情報を出力するシーケンサと、シーケンサの演奏指示に従って楽音を合成する楽音合成装置と、シーケンサの演奏指示に従って歌声を合成する歌声合成装置と、楽音合成装置の出力と歌声合成装置の出力を加算する加算器とを備える。

【0043】

【発明の実施の形態】本発明の請求項1に記載の発明は、入力された音韻と音高と音量および発音タイミングからなる演奏情報を第1および第2の出力系列に交互に出力するアサイナと、前記アサイナの第1および第2の出力系列によってそれぞれ音高と音韻と発音タイミングを指示される第1および第2の読み出し器と、前記アサイナの第1および第2の出力系列によってそれぞれ音量と発音タイミングを指示され波形包絡データを生成する第1および第2のエンベ形成器と、予め複数の音韻および音高の音声波形を記憶していてそれぞれ第1および第2の前記読み出し器によってデータが読み出される第1および第2の波形メモリと、前記第1および第2の読み出し器の出力をそれぞれ前記第1および第2のエンベ形成器の出力でレベル変換する第1および第2のレベル変換器と、前記第1および第2のレベル変換器の出力を加算する加算器とを備えたものであり、指定された音韻/音高の音声波形を、音韻毎に交互に個別に読み出し器で読み出し、クロスフェードしながら滑らかに出力する。クロスフェード部以外は原音PCMそのままなので品質の高い歌声が合成できる。

【0044】また、前の音韻と後ろの音韻で音高差がある場合にはクロスフェード時に前の音韻の音高を次第に後ろの音高に漸近させて音高の変化が滑らかになり、自然な歌声を合成することができる。

【0045】さらに、上記歌声合成装置と楽音合成装置をカラオケに組み込みバックコーラスを発生させて演奏する音程を変化させた時も音声波形の読み出し位置が変化するだけなので演奏速度と音程にかかわらず品質の高いバックコーラスが得られる。

【0046】以下、本発明の実施の形態について、図面を参照しながら説明する。

(実施の形態1) 図1は、本発明の実施の形態1における歌声合成装置の構成を示すブロック図である。図1において、100、101、102、103は入力端子、

104はアサイナ、105a、105bは波形メモリ、106a、106bは読み出し器、107a、107bはエンベ形成器、108a、108bはレベル変換器、109は加算器、110は出力端子、111は音程差分発生器である。

【0047】図2は本実施の形態における読み出し器106a、106bの動作および構成を示す図である。図2(b)において、701、702、703、708は入力端子、707は出力端子、709、706は加算器、704はアドレス生成器、705は読み出し開始アドレス生成器である。

【0048】図3は本実施の形態における音程差分発生器111の動作を説明する説明図、図4は本実施の形態における波形メモリ105aおよび105bのデータ格納状態の説明図、図5は本実施の形態における波形メモリ105aおよび105bのループ読み出しの説明図、図6は本実施の形態における歌声合成のクロスフェード処理の説明図、図7は本実施の形態におけるアサイナ104の動作説明図である。

【0049】以上のように構成された本実施の形態における歌声合成装置の動作を説明する。

【0050】アサイナ104には、合成すべき歌声の音韻と音高と音量および発音タイミングからなる演奏情報がそれぞれ入力端子100、101、102、103から与えられる。アサイナ104は、入力された音韻と音高と音量および発音タイミングからなる演奏情報を交互に第1および第2の出力系列に出力する。

【0051】さらに、アサイナ104の第1および第2の出力系列にはそれぞれ第1の読み出し器106aと第2の読み出し器106bが接続されていて、アサイナ104から音高と音韻と発音タイミングが指示される。

【0052】読み出し器106aおよび106bは次のように動作する。音韻と音高が指示され発音の開始が指示されると、読み出し器106aおよび106bは波形メモリの指示された音韻および音高の格納されている先頭番地から順に波形データを読出す。

【0053】一般に、音声の波形は、発音開始部の数100m秒が過渡部であり、以後は比較的類似した波形が繰り返し出現する性質がある。この繰り返し部分も数1

40 00m秒で表現できることが知られている。すなわち元となる音声の波形を図5(a)とすると、発音開始部+繰り返し部を波形メモリに記憶しておき、読出す際には図5(b)のように繰り返し部をループして読出すことによって、波形メモリ量を1音韻あたり500m秒から1秒程度にすることができる。

【0054】このように波形メモリ105a、105bは図2に示すような構成をしている。入力端子703から音韻、708から音高が指定される。していされた音韻と音高の組み合わせにしたがい読み出し開始アドレス50が読み出し開始アドレス生成器705から出力される。

11

【0055】同時に入力端子701から発音タイミング、702から音程差分値があたえらる。加算器709は音程差分値と音高の指定値を加算しこれを△Aとして出力する。△Aはアドレス生成器704に与えられていてアドレス生成器704は基本的に△Aを積分する。ただし積分値が波形データ長を超える毎に積分値から繰り返し部の長さに相当する値が引き算され、図2(a)のようなアドレス値が生成される。このアドレス値が読み出し開始アドレス生成器705の出力と加算器706で加算され出力端子707に波形メモリに与えるアドレスが出力される。

【0056】また、音程差分発生器111にはアサイナ104の第1および第2の出力系列双方から音高および発音タイミングが指示される。音程差分発生器111は図3のように前の音韻から後ろの音韻に移行する際に前の音韻の音高をしだいに後ろの音韻の音高に近づけるために時間的に変化する第1および第2の補正值をそれぞれ第1および第2の出力とし、これらの出力はそれぞれ第1の読み出し器106aと第2の読み出し器106bの音程変化値入力として与えられる。

【0057】さらに、アサイナ104の第1および第2の出力系列にはそれぞれ波形包絡データを生成する第1のエンベ形成器107aおよび第2のエンベ形成器107bが接続されていてそれぞれアサイナ104の第1および第2の出力系列から音量と発音タイミングを指示される。

【0058】第1の読み出し器106aと第2の読み出し器106bによって読み出される第1の波形メモリ105aおよび105bには図4のように日本語や英語その他の言語について通常発音可能または演奏に必要な全音韻、全音高の波形データが格納されていて、第1の読み出し器106aと第2の読み出し器106bは指定された音高および音韻の波形データを順次読み出す。

【0059】第1の読み出し器106aと第2の読み出し器106bによって読み出された波形データは、それぞれ第1および第2のエンベ形成器107aと107bの出力値に従ってレベル変換器108aと108bによってレベル変換される。

【0060】さらに、レベル変換器108aと108bの出力は加算器109で加算され出力端子110に歌声合成出力が得られる。

【0061】以上の動作の一例として、たとえば、図7の楽譜で示す歌声を合成する場合を説明する。

【0062】この例の場合、アサイナ104の入力および第1、第2の出力系列の出力は図7のようになる。

【0063】アサイナ104の出力は第1の読み出し器106aおよび第2の読み出し器106bに与えられ、第1の読み出し器106aは時間t1から音韻あ(音高ド)を読み出し、時間t3から音韻た(音高ファ)を読み出す。同じく第2の読み出し器106bは時間t2から音韻し

12

(音高ラ)を読み出し、時間t4から音韻の(音高ソ)を読み出す。

【0064】また、アサイナ104の出力は第1のエンベ形成器107aおよび第2のエンベ形成器107bに与えられそれぞれ図6のEaおよびEbに示すような波形エンベロープを発生する。すなわち前後する音韻と音韻の交替部分で、双方の音韻波形をクロスフェードするように波形エンベロープを発生するのである。

【0065】10 このように形成された波形エンベロープEaとEbと読み出し器106aの出力Sa、読み出し器106bの出力Sbがそれぞれレベル変換器108a、108bに与えられる。

【0066】エンベ形成器107aと107bではそれぞれEa×SaとEb×Sbを演算し、エンベ形成器107aと107bの出力が加算器109で加算され歌声合成処理が完了して出力端子110に出力波形が得られる。

【0067】20 同時に音程差分発生器111は以下のように動作する。アサイナ104の出力を受けて、音程差分発生器111は図3のように第1および第2の補正出力を発生する。例えば時間t2で音韻し(音高ラ)の音声を発音開始するときに、直前で音韻あ(音高ド)の音声を発音していて、この間に900セントの音高差がある。この音高差を次第に小さくするために音程差分発生器111は図3に示すように第1の補正出力を第1の読み出し器106aに与える。

【0068】30 前述したように、読み出し器106a、106bは、音高指定値と補正値を加算した値を波形データの読み出し速度に反映するので、音韻あは音高ドから音高ラにしだいに変化することになる。このようにして直前の音韻の音高から次の音韻の音高になめらかに漸近することになり前述したクロスフェード処理との相乗作用により、極めて滑らかに音韻と音韻が接続されて、品質のよい歌声を合成することができる。

【0069】以上のように、本実施の形態によれば、予め複数の音高および音韻の波形データを格納した波形メモリ105aと105bを2つの波形読み出し器106aと106bで交互に読み出し、この読み出し出力をクロスフェードするようにエンベロープを発生するエンベ40 発生器107aと107bを設けこれにより波形読み出し器106aと106bの出力をレベル変換して加算し、さらに前後する音韻の音高が異なる場合には前の音韻から後ろの音韻に移行する際に前の音韻の音高値をしだいに後ろの音高値に近づけるように制御する音程差分値発生器111を設けて、これによって読み出し器106aと106bを制御することにより、極めて滑らかに音韻と音韻が接続することができ、品質のよい歌声を合成することができる。

【0070】50 なお、本実施の形態では、波形メモリ105aと105b、読み出し器106aと106b、エン

ペ形成器107aと107b、レベル変換器108aと108bとしたが、これらを時分割多重処理として単一のハードで構成可能なことはいうまでもない。

【0071】また、レベル変換を乗算としたが、例えばビットシフト等でもかまわない。

(実施の形態2) つぎに本発明の実施の形態2について、図面を参照しながら説明する。

【0072】図8は本発明の実施の形態における音楽再生装置の構成を示すブロック図である。図8において、200は演奏データメモリ、201はシーケンサ、202は歌声合成装置であり、これは実施の形態1で説明したものと同じである。203は楽音合成装置、204は加算器、205は出力端子、206は移調入力端子、207は演奏速度入力端子である。

【0073】以上のように構成された本実施の形態の音楽再生装置について、以下その動作を説明する。

【0074】演奏データメモリ200には演奏情報が格納されている。演奏情報の格納形式は従来例で説明したとおりであり、ここでは繰り返し説明しない。シーケンサ201は演奏データメモリ200から演奏情報を読み出し、歌声の演奏であれば歌声合成装置202に合成すべき歌声の音韻と音高と音量および発音タイミングからなる演奏情報を与え、楽器音の演奏であれば楽音合成装置203に音高と音量および発音タイミングからなる発音情報を与える。

【0075】歌声合成装置202と楽音合成装置203の出力は加算器204で加算され出力端子205に歌声と楽器演奏音が同時に得られる。

【0076】206は移調指示値の入力端子であり、たとえば値1が入力されるとシーケンサ201は音高指示を半音高くして歌声合成装置202および楽音合成装置203に与える。207は演奏速度指示値の入力端子であり、ここに与える値によりシーケンサは演奏速度を速くまたは遅くする。

【0077】本実施の形態では歌声合成装置202に実施の形態1と同じ構成のものを用いているので、入力端子206から移調指示があって、シーケンサ201が音高指示値を変化すると、波形メモリ105a、105bを読み出すアドレスが変化する。

【0078】しかしこれは、読み出す波形データの品質を変化させるものではないので、移調する/しないに係わらず、歌声合成器は劣化するものではない。

【0079】同様に演奏速度指示値が変化しても、読み出し器106aと106b、エンベ形成器107aと107b等に与えるタイミング情報が変化するだけであり、移調の場合と同じく読み出す波形データの品質を変化させるものではないので、移調する/しないに係わらず、歌声合成装置は劣化するものではない。

【0080】以上のように、演奏情報が格納されている演奏データメモリ200と、シーケンサ201と、実施

の形態1による歌声合成装置202と、楽音合成装置203とを設けることにより、演奏速度や音程を変化しても歌声の品質を劣化させることなく音楽を再生することができ、カラオケ等の音楽再生時にキーボードなどの操作をしても、音楽の品質が劣化しない。

【0081】

【発明の効果】以上のように本発明によれば、音韻と音高と音量および発音タイミングからなる演奏情報を第1および第2の出力系列に交互に出力するアサイナと、ア

10 サイナの第1および第2の出力系列によって音高と発音タイミングを指示されて第1および第2の出力にそれぞれ第1および第2の音程変化値を出力する音程差分発生器と、アサイナの第1および第2の出力系列によってそれぞれ音高と音韻と発音タイミングを指示されかつ音程差分発生器の第1および第2の出力によってそれぞれ音程変化値を指示される第1および第2の読み出し器と、アサイナの第1および第2の出力系列によってそれぞれ音量と発音タイミングを指示され波形包絡データを生成する第1および第2のエンベ形成器と、予め複数の音韻および音高の音声波形を記憶していてそれぞれ第1および第2の読み出し器によってデータが読み出される第1および第2の波形メモリと、第1および第2の読み出し器の出力をそれぞれ第1および第2のエンベ形成器の出力でレベル変換する第1および第2のレベル変換器と、第1および第2のレベル変換器の出力を加算する加算器とを備することにより、指定された音韻/音高の音声波形を、音韻毎に交互に個別に読み出し器で読み出し、クロスフェードしながら滑らかに出力する。クロスフェード部以外は原音PCMそのままなので品質の高い歌声が合成でき、さらに前の音韻と後ろの音韻で音高差がある場合にはクロスフェード時に前の音韻の音高を次第に後ろの音高に漸近させて音高の変化が滑らかになり、自然な歌声を合成することができる。

【0082】また、楽器および歌唱の演奏情報を記憶している演奏データメモリと、演奏データメモリから読みだした演奏情報と移調入力および演奏速度入力に応じてタイミングを制御しながら演奏指示情報を出力するシーケンサと、シーケンサの演奏指示に従って楽音を合成する楽音合成装置と、シーケンサの演奏指示に従って歌声40を合成する歌声合成装置と、歌声合成装置の出力と歌声合成装置の出力を加算する加算器とを備することにより、演奏する音程や速度を変化させた時も品質の高い歌声付きの音楽再生ができる、これをシンセカラオケの音楽再生装置に適用することにより、演奏速度や音程が変化してもバックコーラスの品質が低下しないシンセカラオケを提供でき、その実用上優れた効果を有するものである。

【図面の簡単な説明】

【図1】本発明の実施の形態1における歌声合成装置の構成を示すブロック図

50 【図2】同実施の形態における読み出し器の動作および

構成を示す説明図

【図3】同実施の形態における音程差分発生器の動作を説明する説明図

【図4】同実施の形態における波形メモリのデータ格納状態の説明図

【図5】同実施の形態における波形メモリのループ読み出しの説明図

【図6】同実施の形態における歌声合成のクロスフェード処理の説明図

【図7】同実施の形態におけるアサイナの動作説明図

【図8】本発明の実施の形態2における音楽再生装置の構成を示すブロック図

【符号の説明】

104 アサイナ

105a、105b 波形メモリ

106a、106b 読み出し器

107a、107b エンベ形成器

108a、108b レベル変換器

109 加算器

111 音程差分発生器

200 演奏データメモリ

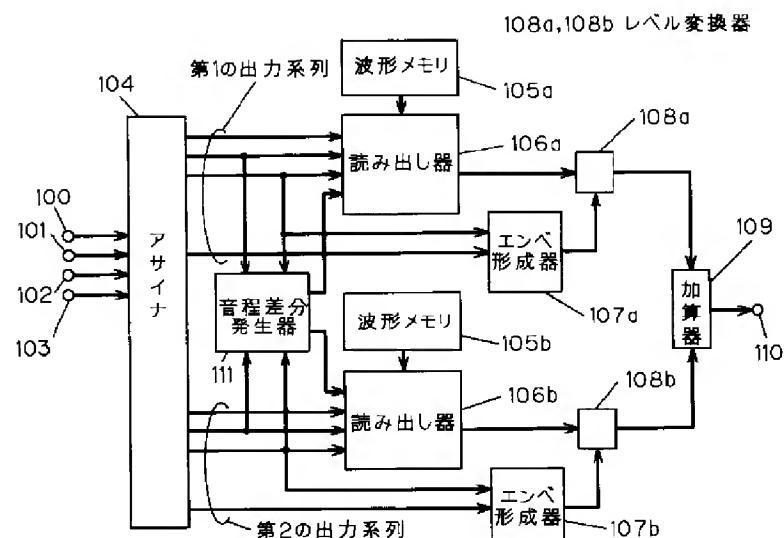
201 シーケンサ

10 202 歌声合成装置

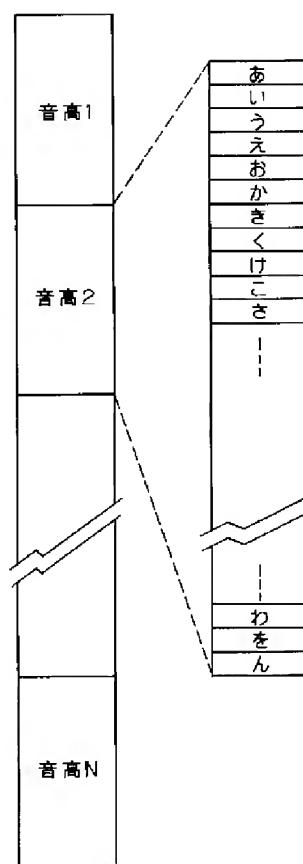
203 楽音合成装置

204 加算器

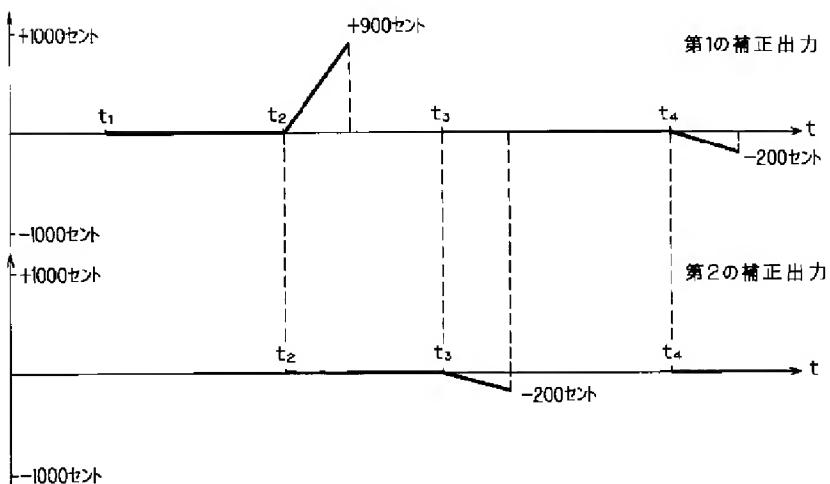
【図1】



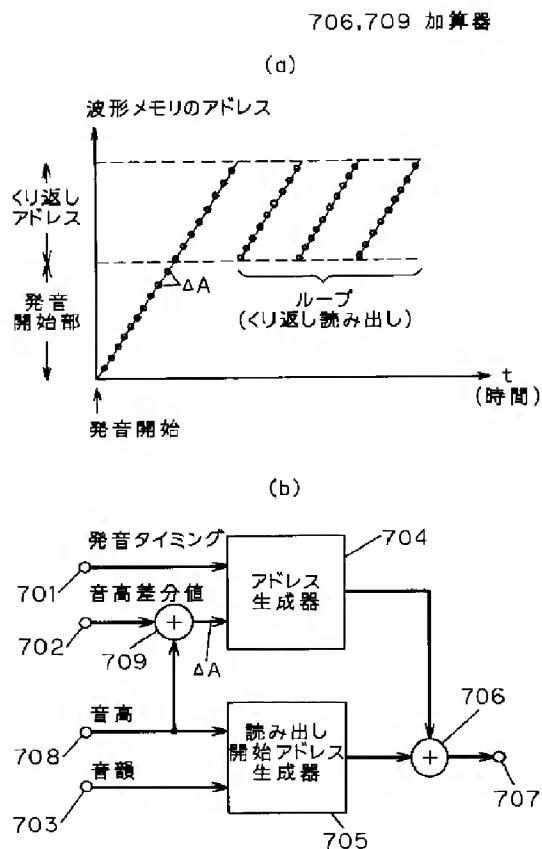
【図4】



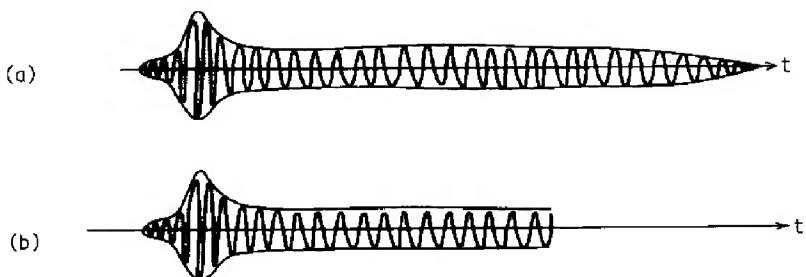
【図3】



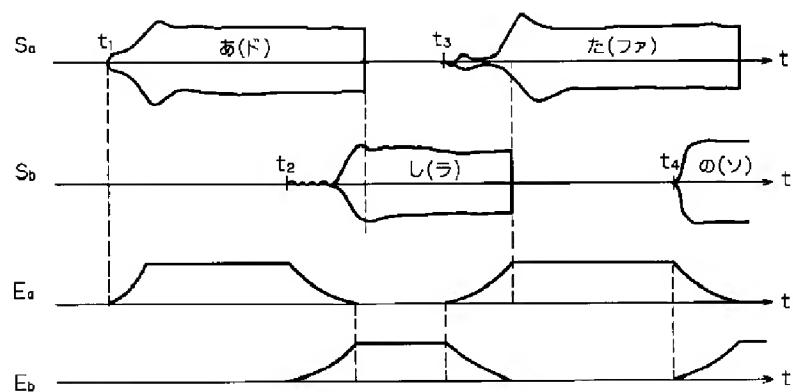
【図2】



【図5】



【図6】



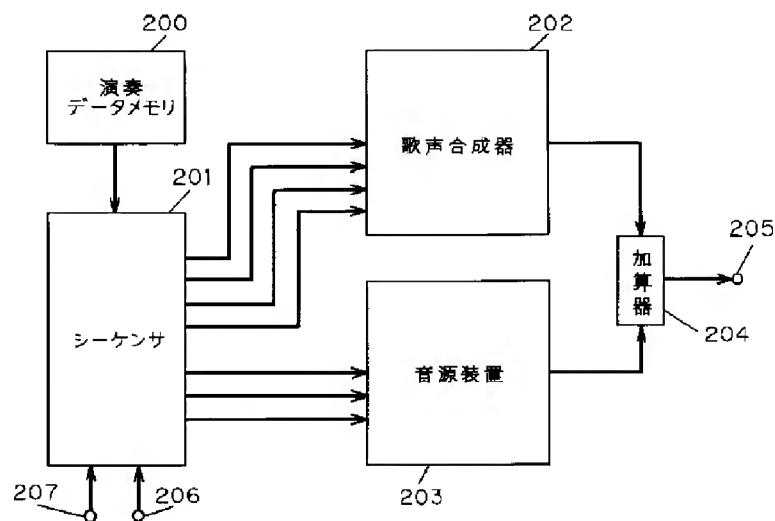
【図7】

アサイン 入力	時間						
	t_1		t_2		t_3		t_4
音韻	あ		し		た		の
音高	ド		ラ		ソ		フア
発音	ON		ON		ON		ON
音量	100		100		100		100

第1の 出力系列	音韻			た		
	あ	し	の	た	ソ	フア
音高	ド	ラ	フア	た	ソ	フア
発音	ON	ON	ON	ON	ON	ON
音量	100	100	100	100	100	100

第2の 出力系列	音韻			の		
	し	の	フア	し	の	フア
音高	ラ	フア	フア	し	の	フア
発音	ON	ON	ON	ON	ON	ON
音量	100	100	100	100	100	100

【図8】



PATENT ABSTRACTS OF JAPAN

(11) Publication number : 09-134188

(43) Date of publication of application : 20.05.1997

(51) Int.Cl. G10K 15/04

G10H 1/00

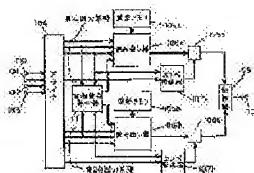
G10H 1/00

G10L 3/02

(21)Application number : 07-292403 (71)Applicant : MATSUSHITA ELECTRIC
IND CO LTD

(22) Date of filing : 10.11.1995 (72) Inventor : HAYASHI KATSUHIKO
MORI DAISUKE

(54) SINGING VOICE SYNTHESIZER AND MUSICAL TONE REPRODUCING DEVICE



(57)Abstract:

PROBLEM TO BE SOLVED: To prevent the degradation in the quality of a back chorus in spite of a change in playing speed and interval by applying a singing voice synthesizer capable of synthesizing the back chorus of PCM quality to a synthesizing KARAOKE machine.

SOLUTION: The waveform data read out by a first reading out device 106a and second reading out device 106b are subjected to level conversion by level converters 108a and 108b according to the output values of first and second envelope formers 107a and 107b. Further, the outputs of these level converters 108a and 108b are added by an adder 109 and the singing voice synthesized output is obtd. at an output terminal 110. Further, the device is provided with an interval differential value generator 111 for controlling the pitch values of phonemes so as to gradually approximate the pitch value of the front phoneme to the rear phoneme at the time the front phoneme transfers to the rear phoneme when the pitches of the front and rear phonemes vary. As a result, the phonemes and the phonemes are extremely smoothly connected by controlling the reading out devices 106a and 106b.

LEGAL STATUS

[Date of request for examination] 10.02.1999

[Date of sending the examiner's decision of rejection] 09.04.2002

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's

decision of rejection]

[Date of requesting appeal against
examiner's decision of rejection]

[Date of extinction of right]

* NOTICES *

**JPO and INPIT are not responsible for any
damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] ASAINA which outputs by turns the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing which were inputted to the 1st and 2nd output sequences, The 1st to which a pitch, a phoneme, and pronunciation timing are directed according to the 1st of said ASAINA, and the 2nd output sequence, respectively, and the 2nd read-out machine, The 1st and the 2nd EMBE formation machine which sound volume and pronunciation timing are directed, respectively and generate wave envelopment data according to the 1st of said ASAINA, and the 2nd output sequence, The 1st and 2nd wave memory from which two or more phonemes and the voice wave of a pitch are beforehand memorized, and data are read with said 1st and 2nd read-out vessels, respectively, The singing voice synthesizer unit equipped with the 1st and 2nd level converters which carry out the level conversion of the output of said 1st and 2nd read-out machines with the output of said 1st and 2nd EMBE formation machines, respectively, and the adder adding

the output of said 1st and 2nd level converters.

[Claim 2] ASAINA which outputs by turns the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing which were inputted to the 1st and 2nd output sequences, the musical interval which a pitch and pronunciation timing are directed and outputs the 1st and 2nd musical interval change values to the 1st and 2nd outputs according to the 1st of said ASAINA, and the 2nd output sequence, respectively -- difference -- with a generator A pitch, a phoneme, and pronunciation timing are directed according to the 1st of said ASAINA, and the 2nd output sequence, respectively. and said musical interval -- difference -- with the 1st [to which a musical interval change value is directed with the 1st and 2nd outputs of a generator, respectively], and 2nd read-out machines The 1st and the 2nd EMBE formation machine which sound volume and pronunciation timing are directed, respectively and generate wave envelopment data according to the 1st of said ASAINA, and the 2nd output sequence, The 1st and 2nd wave memory from which two or more phonemes and the voice wave of a pitch are beforehand memorized, and data are read with said 1st and 2nd read-out vessels, respectively, The singing voice synthesizer unit equipped with the 1st and 2nd level converters which carry out the level conversion of the output of said 1st and 2nd read-out machines with the output of said 1st and 2nd EMBE formation machines, respectively, and the adder adding the output of said 1st and 2nd level converters.

[Claim 3] The performance data memory which has memorized a musical instrument and the performance information on a song, and the sequencer which outputs performance directions information while controlling timing according to the performance information, transposition input, and performance rate input which were read from said performance data memory, The musical-sound synthesizer unit which compounds musical sound according to performance directions of said sequencer, the singing voice synthesizer unit which compounds singing voice according to performance directions of said sequencer, and the music regenerative apparatus equipped with the adder adding the output of said

musical-sound synthesizer unit, and the output of said singing voice synthesizer unit.

[Claim 4] A singing voice synthesizer unit is a music regenerative apparatus according to claim 3 characterized by being a configuration according to claim 1 or 2.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
 2. **** shows the word which can not be translated.
 3. In the drawings, any words are not translated.
-

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the singing voice synthesizer unit and music regenerative apparatus which do not fall the quality of a back back chorus by II Tempo (performance rate) of music, and pitch (transposition) modification especially about the music regenerative apparatus used for the karaoke equipment which performs the karaoke performance with a back chorus.

[0002]

[Description of the Prior Art] In recent years, with an advance of digital technique, various kinds of karaoke methods are proposed and the karaoke (synthesizer karaoke is called hereafter) of the synthesizer method using the same musical-

sound synthesizer unit especially as electrohone attracts attention increasingly.

[0003] Synthesizer karaoke drives a musical-sound synthesizer unit using the performance directions information generally called MIDI, and generates the back music for karaoke, and the features are in that there is very little amount of data which the performance of one music takes compared with general karaoke (for example, karaoke of a laser disc).

[0004] These features of a synthesizer method are efficiently employed in the synthesizer karaoke of a type which stores in semiconductor memory etc. the synthesizer karaoke (the so-called online karaoke) which transmits ISDN and the telephone line for performance information as a transmission line, and performance information, and growth of the synthesizer karaoke in this field will be expected from now on.

[0005] On the other hand, the fault that the performance of voice or a back chorus is difficult exists in the karaoke of a synthesizer method. synthesizer karaoke is constituted using the musical-sound synthesizer unit of electrohone, and, as for this, the musical-sound synthesizer unit is constituted for the purpose of playback of a musical instrument sound (musical instruments other than voice, such as a piano and a guitar) -- the thing reason is carried out.

[0006] There are some which were indicated by JP,6-161479,A as such a music regenerative apparatus of the karaoke of the synthesizer method conventionally devised in order to conquer a fault.

[0007] That explanation is hereafter given for this conventional music regenerative apparatus using a drawing given [this] in an official report. drawing 1 of this official report -- setting -- 1 -- the musical-sound data-logging section and 2 -- for the alphabetic character data-logging section and 5, as for a synthetic amplifier and 7, the speech synthesis section and 6 are [the musical-sound performance section and 3 / the input section and 4 / the output section and 8] synchronizers.

[0008] Explanation is given below about the conventional music regenerative apparatus constituted as mentioned above. The musical-sound data-logging

section 1 records the sound recording data of performance music, and consists of a compact disk (CD), a laser disc (LD) or a magnetic-recording medium, and its drive equipment. The musical-sound performance section 2 changes the musical-sound data of the musical-sound data-logging section 1 into a receipt and the musical-sound signal of an analog. What is necessary is just to use the synthesizer by which the MIDI code, then the musical-sound performance section 2 generate a sound for the musical-sound data currently recorded on the Records Department 1 here using an electronic oscillator. Moreover, the input section 3 inputs a song person's voice with a microphone etc. Moreover, it is also possible for the alphabetic character data-logging section 4 to record a character code train, and for drive equipment to be consisted of by a magnetic-recording medium etc. and ** like the musical-sound data-logging section 1, and to serve with the musical-sound data-logging section 1. The speech synthesis section 5 changes into a receipt sound signal the alphabetic character data stream recorded on the alphabetic character data-logging section 4. Moreover, the synthetic amplifier 6 compounds and amplifies the output of the sound signal inputted from the output of the musical-sound performance section 2, and the input section 3, and the speech synthesis section 5. The output of the result is sent to the output sections 7, such as a loudspeaker. Moreover, a synchronizer 8 orders it control of the alphabetic data sent to the speech synthesis section 5 from the alphabetic character data-logging section 4 according to the performance of the musical-sound performance section 2.

[0009] The MIDI code shall be recorded on the music data-logging section 1. Drawing 2 of this official report shows the example of the alphabetic data currently recorded on the alphabetic character data-logging section 4. Actuation of this equipment is explained with reference to drawing 1 and drawing 2 . If the MIDI code begins to be performed, the value of the readout rate from the musical-sound data-logging section 1 set as the musical-sound performance section 2, i.e., a musical-sound performance rate, will be sent to a synchronizer 8, and will be memorized to the synchronizer memory 9. Next, the data of the 1st

vibrant tune of the MIDI code are read from the musical-sound data-logging section 1, are sent to the musical-sound performance section 2, and are changed into a musical-sound signal. Having received the MIDI code for one vibrant tune to coincidence is told to a synchronizer 8. If it gets to know that the MIDI code of one vibrant tune was sent to the musical-sound performance section 2, this synchronizer 8 will read one-line data from the alphabetic character data-logging section 4, will count the number of alphabetic characters, and will calculate the value of audio speed assignment. Here, the alphabetic data for one vibrant tune is made into one line to line feed. And the speech synthesis section 5 sends the character code data pronounced from assignment of speed. The value of speed is calculated as follows. Value = (time amount of one vibrant tune)/of speed (the number of alphabetic characters of one line).

[0010] For example, it supposes that it was set up so that the music which it is going to perform now might be performed with the speed for 1 vibrant-tune 5 seconds, and supposing the number of alphabetic characters of one line of a back chorus is ten characters, the speech synthesis section 5 will change the alphabetic character sent with the specified speed into a sound signal, and will send it to the synthetic amplifier 6. Consequently, the synchronization with the musical sound generated in the musical-sound performance section 2 and the voice generated in the speech synthesis section 5 is taken, and it is sent to the output section 7 from the synthetic amplifier 6. Thus, since vocal sound data are recorded as a character code train and he is trying to generate voice from the character code train, vocal sound data as well as the MIDI data of musical sound can be coded, and can be recorded, and data can be compressed sharply.

[0011] However, although it can direct that only the alphabetic character is memorized in the music regenerative apparatus indicated here as shown in the alphabetic character data-logging section 4 at drawing 2 , and these alphabetic characters are changed into a phoneme in the speech synthesis section 6 In being unable to direct a musical interval to the speech synthesis section 6, since the pronunciation timing of these phonemes was also equally assigned within 1

vibrant tune, there was finely uncontrollable constraint, and the technical problem that only the special back chorus for musical pieces was realizable occurred.

[0012] There is karaoke equipment indicated by JP,7-140991,A as a music regenerative apparatus of the conventional synthesizer karaoke which improved this point.

[0013] Hereafter, the explanation is performed for the conventional music regenerative apparatus to below using the drawing of this official report.

[0014] drawing 1 of this official report -- setting -- 1 -- CPU and 2 -- RAM and 3 -- a bus and 4 -- music data storage and 5 -- panel I/F and 6 -- a controller and 7 -- background image storage / playback section and 8 -- an image / words display, and 9 -- a video selector and 10 -- for a mixer and an effector, and 13, as for a sound source and 15, amplifier and a loudspeaker, and 14 are [a monitor and 11 / a microphone and 12 / Program ROM and a sequencer, and 16] digital voice decoders.

[0015] Drawing 2 of this official report is structural drawing of music data, drawing 3 is the explanatory view showing the detail of performance data tracks, drawing 4 is the explanatory view showing the detail of a voice data directions truck, drawing 5 is serial structural drawing of each truck, drawing 6 is the explanatory view of the part which decodes voice data, and drawing 7 is the timing diagram of a chorus voice control.

[0016] The actuation is explained below about the conventional karaoke equipment constituted as mentioned above.

[0017] Drawing 1 is the block diagram of the whole karaoke equipment of the conventional example. RAM (random access memory) used when CPU (central processing unit) to which it sets to this drawing, and 1 controls and manages system-wide actuation, and 2 control and manage system-wide actuation by CPU1, and 3 are the data and the address buses for unifying the whole system.

[0018] The store with which 4 stored two or more music data, for example, HDD, (hard disk) Controllers, such as two or more remote control which 5 minds panel I/F (interface), and 6 minds panel I/F5, and gives directions to a system, An

image / words display for 7 to display background image storage / playback section, and for 8 display the still picture and words of a background, The video selector which 9 chooses the background image (animation) from background image storage / playback section 7 and the image from an image / words display 8, and is compounded, and 10 are monitors which display the image which was chosen by the video selector and compounded.

[0019] The microphone into which 11 inputs a singer's song voice, and 12 are effectors which give the mixer and various kinds of sound effects which mix song voice and the musical sound of performance music. 13 is a loudspeaker which amplifies and outputs with amplifier the song voice and the performance music which were compounded. The sound source in which 14 pronounces the musical sound of performance music, and 15 are sequencers which control a mixer and an effector 12 in this sound-source 14 list. This sequencer 15 has the program ROM which stored the program which CPU1 uses. 16 is a voice decoder which decodes the encoded digital voice data (for example, PCM and ADPCM data).

[0020] Hereafter, actuation is explained. If performance music is specified by actuation of a controller 6, CPU1 will transmit the voice data with which it encoded the music data of the performance music which corresponds with reference to the music list memorized by the store 4, and for back choruses to RAM2, and will move control to a sequencer 15. A sequencer 515 arranges in parallel and performs two or more events including a music performance to coincidence based on two or more event data contained in music data.

[0021] That is, further, a background image number is supplied to background image record / playback section 7, and the voice data with which data concerning [a sequencer 15] the tone of a musical piece and the tone of false voice were encoded by the sound source 14 again supplies a words number to the digital voice decoder 16 at an image / words display 8, respectively. Consequently, a background image is displayed all over a monitor 10, and while it will be in the display condition which words superimposed to that part, it will be in the performance condition that performance music and a back chorus are outputted

from a loudspeaker 13.

[0022] Music data consist of a header unit, the data sequence section, and the voice data section, as shown in drawing 2 . A tune number number, a music name, a composer name, a singer name, background-image selection information, and the font information on words are written in the header unit as information on the music proper.

[0023] Two or more trucks which described two or more kinds of events performed by coincidence in parallel are set to the data sequence section. Two or more voice numbers chosen by the voice data directions data of the data sequence section are written in the voice data section.

[0024] The data which make performance data tracks utter the musical sound of performance music from a sound source 14 among the trucks included in the data sequence section are described to time series. the data which make false voice data tracks utter false chorus voice (for example, "***- -- -- obtaining -") from a sound source 14 are described to time series. in the voice data directions truck, each data of the voice data number and musical interval which direct the class of intrinsic chorus voice (for example, it is "**** -- **- and "**** point -") made to decode by the decoder 16, and a pitch is described. The data which direct the class of words which combine with a performance and are displayed on a monitor 10 in words data tracks are described to time series. CDC which controls a mixer and an effector 12 in effectiveness control data tracks is described to time series.

[0025] Drawing 3 is the explanatory view showing the detail of - ** truck by performance. Each information on a note event, a tone modification event, and a pitch bend event is described by these performance data tracks. CH number which directs one CH (channel) which a sound source 14 tends to make it utter, a note number (pitch), a velocity (sound volume), and ** length are written in the note event. CH number and tone data are written in the tone modification event. CH number and pitch bend information are written in the pitch bend event. Although the note event of performance data tracks is an object for the musical-

sound pronunciation of performance music, it has the note event of structure only similarly [false voice data tracks] from which CH number differs.

[0026] Drawing 4 is the explanatory view showing the detail of a voice data directions truck. Each information on a voice directions event is described by this voice data directions truck. That is, it is each information on a voice data number, a musical interval, and sound volume. A voice data number is a number of the encoded intrinsic chorus voice data which is made to decode by the decoder 16, and is a voice number of the voice data of drawing 2 . Beforehand, two or more trucks are the class of event, and the structure which arranged latency-time deltat to the next event utterance to time series, as shown in drawing 5 .

[0027] Drawing 6 is the explanatory view of the part which decodes voice data according to the information on a voice directions event. A sequencer 15 reads the digital voice data with which the voice data number which corresponds from RAM2 using the voice data number of a voice directions event was encoded, and inputs it into the digital voice decoder 16. This digital voice data is intrinsic chorus voice data in the conventional example. When digital voice data is adaptive (delta AD) PCM data compressed in order to reduce data volume as an example, a decoder 16 is a PCM decoder with the function which performs number-of-bits conversion and frequency conversion, and is elongated.

[0028] In the latter part of a decoder 16, the processor 17 which controls a musical interval and sound volume is arranged, and it inputs into a mixer and an effector 12, after adjusting the musical interval and sound volume of an analog wave which were decoded here based on the musical interval and sound-volume information which are included in a voice directions event. When reproduction speed is 100%, the time of the amount of modification of a pitch being predetermined within the limits reproduces intrinsic chorus voice using this path. When the reproduction speed of performance music is changed with the directions from a controller 6 in addition to 100%, in a processor 17, sound volume to the output of a decoder 16 is set to 0, and playback of intrinsic chorus voice is stopped. It is also the same as when the amount of modification of a

pitch exceeds 200-300 cents of the specified quantity.

[0029] Not using intrinsic chorus voice at the time of pitch modification is based on the following reason. When the musical interval to intrinsic chorus voice is changed to some extent above, the audio quality deteriorates and it becomes impossible that is, to use it substantially in a musical interval and the sound-volume control section 17. "-- to some extent --" -- a standard -- semitone -- it is **2 - 3 sound (200 to 300 cents) in a unit. Therefore, not only when II Tempo (reproduction speed) is changed, but when it changes a pitch (musical interval) more than the specified quantity, false chorus voice is used.

[0030] When the reproduction speed of performance music consists of the above reason in addition to 100%, or when the amount of modification of a pitch exceeds 200-300 cents of the specified quantity, a sequencer 515 makes a false chorus utter from a sound source 514 according to the data of false voice data tracks. false chorus voice -- beforehand -- a sound source 514 -- musical sound -- the same -- carrying out -- "***" -- "it obtains, the channel of the voice -" is set up and it is made to choose and pronounce with false voice data

[0031] Drawing 7 shows the situation of control by the sequencer 15 when II Tempo (reproduction speed) changes focusing on 100%. In this example, the event of voice data tracks changes in order at time of day t1, t5, and t8. On the other hand, II Tempo is the example which changed to 120% from 70% at time of day t6, and changed [% / 1005 to 70] from 120% to 100% at time of day t7 by time amount t4.

[0032] since II Tempo of time of day t1 is 100%, according to the event of voice data tracks, it is "****" at this time -- decoding of the intrinsic voice data **- is started and the sound volume of voice data is set as the specified quantity described by the voice event. Nothing is not necessarily processed about false voice data tracks, the voice event equivalent to the first "***" is read, and that false voice (for example, "**-") is made to pronounce from a sound source 14 at this time. However, sound volume by the side of false voice is set to 0, and it prevents from outputting from a loudspeaker 13 in fact. To coincidence, the

sound-volume value (the above-mentioned predetermined value) about intrinsic voice data is memorized for the return.

[0033] Since time of day t2 is also 100%, II Tempo continues the output of intrinsic chorus voice for it. Although a false voice truck goes into the pronunciation event of "***" at this time, the sound volume by the side of false voice is still 0. Since time of day t3 is also 100%, II Tempo continues the output of the voice of an intrinsic chorus for it. this time -- a false voice truck -- "-- it is --" -- although it goes into a pronunciation event, the sound volume by the side of false voice is still 0.

[0034] Since II Tempo will fall from 100% to 70% if time of day t4 comes, fade-out is carried out [voice / intrinsic] and fade-in is carried out to the predetermined value which memorized the pronunciation channel of false voice beforehand instead. At this time, since intrinsic voice may be in the middle of playback, loose cross fade is performed and a way piece is avoided.

[0035] having set sound volume to 0, since II Tempo had not returned to 100% although the voice data side went into the following event ("***** -- things -- ") when time of day t5 came -- the pronunciation of intrinsic voice -- not carrying out -- instead of -- " -- carrying out -- " -- the receiving false voice (for example, "***-") is made to utter from a sound source 14 About this false voice, it controls using CH number described by the event, a note number (pitch), a velocity (sound volume), and ** length.

[0036] Although II Tempo changes to 120% from 70% at time of day t6, since it is not 100%, the pronunciation of the false voice (for example, "***-") equivalent to "***" is continued. Since II Tempo changed to 100% from 70% at time of day t7, one condition which returns to intrinsic voice was satisfied, but the pronunciation of false voice is continued until the following voice data event begins, since it is in this phase in the middle of decoding of a front voice data event.

[0037] Since the new voice data event whose II Tempo is 100% will begin if time of day t8 comes, fade-in is carried out [voice / intrinsic] to this timing. Since it is at the initiation time of intrinsic voice at this time, it is satisfactory even if it carries

out fade-in steeply. Although there is also a pronunciation part equivalent to "****" of false voice at this time, sound volume of that pronunciation channel is set to 0, and pronunciation is forbidden. Even if reproduction speed changes karaoke performance music with chorus PERT because a sequencer 15 performs the above control, it can provide in the condition that there is always no time lag.

[0038]

[Problem(s) to be Solved by the Invention] However, since it would change from the back chorus formed in PCM and the PCM compressional-wave form to the false back chorus formed with the musical-sound synthesizer unit if a performance rate is changed or a musical interval is changed with the above-mentioned conventional configuration, tone quality changed, and when especially a performance rate and a musical interval were not in default value, it became a false back chorus, and had the technical problem that the quality of a back chorus deteriorated. The false back chorus constituted from "****" and "****" to the back chorus formed especially by PCM and PCM compression reproducing words is heterogeneous.

[0039] Moreover, the amount of data of the back chorus of PCM and PCM compressed format takes great time amount to transmit these via a communication line for every music in online karaoke, since many [very] (100 or more times) compared with the performance information which drives a musical-sound synthesizer unit and which is generally called MIDI, and a lot of data storage is needed in the synthesizer karaoke of a stand-alone. In order to avoid this, even if the performance rate or the musical interval was in default value, it had the technical problem that it had to be made to have to use a false back chorus.

[0040] This invention offers the music regenerative apparatus with which the quality of a back chorus does not deteriorate even if a performance rate and a musical interval change by solving the above-mentioned conventional technical problem, offering the singing voice synthesizer unit which can compound the back chorus of PCM quality, and applying this to synthesizer karaoke further.

[0041]

[Means for Solving the Problem] In order to solve these technical problems, the singing voice synthesizer unit of this invention ASAINA which outputs by turns the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing which were inputted to the 1st and 2nd output sequences, the musical interval which 4s pitches and pronunciation timing are directed by the 1st of ASAINA, and the 2nd output sequence, and outputs the 1st and 2nd musical interval change values to the 1st and 2nd outputs, respectively -- difference -- with a generator a pitch, a phoneme, and pronunciation timing are directed according to the 1st of ASAINA, and the 2nd output sequence, respectively -- having -- and a musical interval -- difference -- with the 1st [to which a musical interval change value is directed with the 1st and 2nd outputs of a generator respectively], and 2nd read-out machines The 1st and the 2nd EMBE formation machine which sound volume and pronunciation timing are directed, respectively and generate wave envelopment data according to the 1st of ASAINA, and the 2nd output sequence, The 1st and 2nd wave memory from which two or more phonemes and the voice wave of a pitch are beforehand memorized, and data are read with the 1st and 2nd read-out vessels, respectively, It has the 1st and 2nd level converters which carry out the level conversion of the output of the 1st and 2nd read-out machines with the output of the 1st and 2nd EMBE formation machines, respectively, and an adder adding the output of the 1st and 2nd level converters.

[0042] Moreover, the performance data memory the music regenerative apparatus of this invention has remembered a musical instrument and the performance information on a song to be, The sequencer which outputs performance directions information while controlling timing according to the performance information, transposition input, and performance rate input which were read from performance data memory, It has the musical-sound synthesizer unit which compounds musical sound according to performance directions of a sequencer, the singing voice synthesizer unit which compounds singing voice

according to performance directions of a sequencer, and an adder adding the output of a musical-sound synthesizer unit, and the output of a singing voice synthesizer unit.

[0043]

[Embodiment of the Invention] ASAINA which outputs by turns the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing of having inputted invention of this invention according to claim 1 to the 1st and 2nd output sequences, The 1st to which a pitch, a phoneme, and pronunciation timing are directed according to the 1st of said ASAINA, and the 2nd output sequence, respectively, and the 2nd read-out machine, The 1st and the 2nd EMBE formation machine which sound volume and pronunciation timing are directed, respectively and generate wave envelopment data according to the 1st of said ASAINA, and the 2nd output sequence, The 1st and 2nd wave memory from which two or more phonemes and the voice wave of a pitch are beforehand memorized, and data are read with said 1st and 2nd read-out vessels, respectively, The 1st and 2nd level converters which carry out the level conversion of the output of said 1st and 2nd read-out machines with the output of said 1st and 2nd EMBE formation machines, respectively, It has an adder adding the output of said 1st and 2nd level converters, and it reads according to an individual by turns for every phoneme, and the voice wave of the specified phoneme/pitch is outputted smoothly, reading and carrying out cross fade with a vessel. except for the cross fade section -- a fundamental tone PCM -- since it remains as it is, the high singing voice of quality is compoundable.

[0044] Moreover, since it carries out asymptotic [of the pitch of a front phoneme] to a back pitch gradually at the time of cross fade when there is a pitch difference by the front phoneme and the back phoneme, change of a pitch becomes smooth, and natural singing voice can be compounded.

[0045] Furthermore, since the read-out location of a voice wave only changes also when changing the musical interval performed since the above-mentioned

singing voice synthesizer unit and a musical-sound synthesizer unit are built into karaoke and a back chorus is generated, the high back chorus of quality is obtained irrespective of a performance rate and a musical interval.

[0046] Hereafter, the gestalt of operation of this invention is explained, referring to a drawing.

(Gestalt 1 of operation) Drawing 1 is the block diagram showing the configuration of the singing voice synthesizer unit in the gestalt 1 of operation of this invention. drawing 1 -- setting -- 100, 101, 102, and 103 -- an input terminal and 104 -- ASAINA, and 105a and 105b -- wave memory, and 106a and 106b -- a read-out machine, and 107a and 107b -- an EMBE formation machine, and 108a and 108b -- a level converter and 109 -- an adder and 110 -- an output terminal and 111 -- a musical interval -- difference -- it is a generator.

[0047] Drawing 2 is drawing showing the actuation and the configuration of the read-out machines 106a and 106b in the gestalt of this operation. For an input terminal and 707, as for an adder and 704, in drawing 2 (b), an output terminal, and 709 and 706 are [701, 702, 703, and 708 / an address-generation machine and 705] read-out starting address generation machines.

[0048] a musical interval [in / in drawing 3 / the gestalt of this operation] -- difference -- the explanatory view explaining actuation of a generator 111, the explanatory view of the data-storage condition of the wave memory [in / in drawing 4 / the gestalt of this operation] 105a and 105b, the explanatory view of loop-formation read-out of the wave memory [in / in drawing 5 / the gestalt of this operation] 105a and 105b, the explanatory view of cross-fade processing of singing-voice composition [in / in drawing 6 / the gestalt of this operation], and drawing 7 are the explanatory views of ASAINA 104 in the gestalt of this operation of operation.

[0049] Actuation of the singing voice synthesizer unit in the gestalt of this operation constituted as mentioned above is explained.

[0050] The performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing of singing voice which should be

compounded is given to ASAINA 104 from input terminals 100, 101, 102, and 103, respectively. ASAINA 104 outputs by turns the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing which were inputted to the 1st and 2nd output sequences.

[0051] Furthermore, 1st read-out machine 106a and 2nd read-out machine 106b are connected to the 1st of ASAINA 104, and the 2nd output sequence, respectively, and a pitch, a phoneme, and pronunciation timing are directed from ASAINA 104.

[0052] The read-out machines 106a and 106b operate as follows. If a phoneme and a pitch are directed and initiation of pronunciation is directed, the read-out machines 106a and 106b will read a data point sequentially from the head address with which the phoneme wave memory was instructed to be, and the pitch are stored.

[0053] Generally, several 100m second of the pronunciation initiation section of an audio wave is the transient section, and there is a property in which a comparatively similar wave appears repeatedly henceforth. It is known that this repeat part can also be expressed in several 100m second. That is, if the audio wave which becomes origin is made into drawing 5 (a), in case the pronunciation initiation section + repeat section is memorized in wave memory and it reads, the amount of wave memory can be made into about 1 second from 500 m seconds per one phoneme by carrying out the loop formation of the section and reading it repeatedly, like drawing 5 (b).

[0054] Thus, the wave memory 105a and 105b is carrying out the configuration as shown in drawing 2 . A pitch is specified from a phoneme and 708 from an input terminal 703. A read-out starting address reads according to the combination of the phoneme which volunteered by carrying out, and a pitch, and it is outputted from the starting address generation machine 705.

[0055] coincidence -- the pronunciation timing from an input terminal 701, and the musical interval from 702 -- difference -- a value -- giving -- ****. an adder 709 -- a musical interval -- difference -- a value and the assignment value of a pitch are

added and this is outputted as deltaA. deltaA is given to the address-generation machine 704 and the address-generation machine 704 integrates with deltaA fundamentally. However, whenever an integral value exceeds data-point length, the value which is repeatedly equivalent to the die length of the section subtracts from an integral value, and an address value like drawing 2 (a) is generated. The address which this address value reads, and it is added with the output and adder 706 of the starting address generation machine 705, and is given to wave memory at an output terminal 707 is outputted.

[0056] moreover, a musical interval -- difference -- a pitch and pronunciation timing are directed in a generator 111 from both 1st [of ASAINA 104], and 2nd output sequences. a musical interval -- difference -- a generator 111 considers the 1st and 2nd correction value which changes in time as the 1st and 2nd outputs, respectively, in order to bring the pitch of a front phoneme close to the pitch of a back phoneme gradually, in case it shifts to a back phoneme from a front phoneme like drawing 3 , and these outputs are given as musical interval change markup force of 1st read-out machine 106a and 2nd read-out machine 106b, respectively.

[0057] Furthermore, 1st EMBE formation machine 107a and 2nd EMBE formation machine 107b which generate wave envelopment data, respectively are connected to the 1st of ASAINA 104, and the 2nd output sequence, and sound volume and pronunciation timing are directed from the 1st of ASAINA 104, and the 2nd output sequence, respectively.

[0058] The data point of all the phonemes [language / Japanese or / of English and others] required for a performance usually that can be pronounced, or whole tone quantity is stored in the 1st wave memory 105a and 105b read by 1st read-out machine 106a and 2nd read-out machine 106b like drawing 4 , and 1st read-out machine 106a and 2nd read-out machine 106b read the data point of the specified pitch and a phoneme one by one.

[0059] According to the output value of the 1st and 2nd EMBE formation machines 107a and 107b, the level conversion of the data point read by 1st read-

out machine 106a and 2nd read-out machine 106b is carried out by level converters 108a and 108b, respectively.

[0060] Furthermore, the output of level converters 108a and 108b is added with an adder 109, and a singing voice composition output is obtained by the output terminal 110.

[0061] The case where the singing voice shown in the score of drawing 7 is compounded as an example of the above actuation is explained.

[0062] In the case of this example, the output of the input [of ASAINA 104] and 1st, and 2nd output sequence becomes like drawing 7 .

[0063] The output of ASAINA 104 is given to 1st read-out machine 106a and 2nd read-out machine 106b, and 1st read-out machine 106a reads ***** (pitch DO) from time amount t1, and reads ***** (pitch FA) from time amount t3. the same -- 2nd read-out machine 106b -- the phoneme from time amount t2 -- carrying out (pitch RA) -- it reads and (pitch SO) of a phoneme is read from time amount t4.

[0064] Moreover, the output of ASAINA 104 generates a wave envelope as given to 1st EMBE formation machine 107a and 2nd EMBE formation machine 107b and shown in Ea and Eb of drawing 6 , respectively. That is, it is the shift parts of a phoneme and a phoneme which get mixed up, and a wave envelope is generated so that cross fade of both phoneme waves may be carried out.

[0065] Thus, it reads with the formed wave envelopes Ea and Eb, and the output Sa of vessel 106a and the output Sb of read-out machine 106b are given to level converters 108a and 108b, respectively.

[0066] With the EMBE formation vessels 107a and 107b, EaxSa and EaxSb are calculated, respectively, the output of the EMBE formation machines 107a and 107b is added with an adder 109, singing voice composition processing is completed, and an output wave is acquired by the output terminal 110.

[0067] coincidence -- a musical interval -- difference -- a generator 111 operates as follows. the output of ASAINA 104 -- winning popularity -- a musical interval -- difference -- a generator 111 generates the 1st and 2nd amendment outputs like drawing 3 . For example, when carrying out [voice / which carries out a phoneme

by time amount t2 (pitch RA)] pronunciation initiation, the voice of ***** (pitch DO) is pronounced just before, and there is a pitch difference of 900 cents in the meantime. in order to make this pitch difference small gradually -- a musical interval -- difference -- a generator 111 gives the 1st amendment output to 1st read-out machine 106a, as shown in drawing 3 .

[0068] As mentioned above, since the read-out machines 106a and 106b reflect the value adding a pitch assignment value and correction value in the read-out rate of a data point, ***** will change from pitch DO to pitch RA gradually. Thus, according to a synergism with the cross fade processing which will carry out [processing] asymptotic to the pitch of the following phoneme smoothly, and was mentioned above from the pitch of the last phoneme, a phoneme and a phoneme are connected very smoothly, and quality singing voice can be compounded.

[0069] As mentioned above, according to the gestalt of this operation, the wave memory 105a and 105b which stored two or more pitches and the data point of a phoneme beforehand is read by turns with two wave read-out vessels 106a and 106b. Form the EMBE generators 107a and 107b which generate an envelope so that cross fade of this read-out output may be carried out, and by this, carry out the level conversion of the output of the wave read-out machines 106a and 106b, and it is added. The value generator 111 is formed. the musical interval controlled to bring the sound high price of a front phoneme close to a back sound high price gradually in case it shifts to a back phoneme from a front phoneme, when the pitches of the phoneme which furthermore gets mixed up differ -- difference -- By reading by this and controlling Vessels 106a and 106b, a phoneme and a phoneme can be connected very smoothly and quality singing voice can be compounded.

[0070] In addition, although considered as the wave memory 105a and 105b, the read-out machines 106a and 106b, the EMBE formation machines 107a and 107b, and level converters 108a and 108b with the gestalt of this operation, in these, it is hard and it cannot be overemphasized that single as time division

multiple processing it can constitute.

[0071] Moreover, although the level conversion was made into multiplication, a bit shift etc. is sufficient, for example.

(Gestalt 2 of operation) The gestalt 2 of the operation of this invention to the next is explained, referring to a drawing.

[0072] Drawing 8 is the block diagram showing the configuration of the music regenerative apparatus in the gestalt of operation of this invention. In drawing 8 , as for performance data memory and 201, 200 is [a sequencer and 202] singing voice synthesizer units, and this is the same as what was explained with the gestalt 1 of operation. For a musical-sound synthesizer unit and 204, as for an output terminal and 206, an adder and 205 are [203 / a transposition input terminal and 207] performance rate input terminals.

[0073] About the music regenerative apparatus of the gestalt of this operation constituted as mentioned above, the actuation is explained below.

[0074] Performance information is stored in the performance data memory 200. the storing format of performance information gets that the conventional example explained, comes out, and there is, and here does not explain it repeatedly. A sequencer 201 reads performance information from the performance data memory 200, if it is the performance of singing voice, it will give the performance information which consists of the phoneme, the pitch, sound volume, and pronunciation timing of singing voice which should be compounded to the singing voice synthesizer unit 202, and if it is the performance of a musical instrument sound, it will receive the pronunciation information which becomes the musical-sound synthesizer unit 203 from a pitch, sound volume, and pronunciation timing.

[0075] The output of the singing voice synthesizer unit 202 and the musical-sound synthesizer unit 203 is added with an adder 204, and singing voice and a musical instrument performance sound are obtained by the output terminal 205 at coincidence.

[0076] if 206 is the input terminal of transposition indicated value, for example, a value 1 is inputted -- a sequencer 201 -- pitch directions -- semitone -- it is made

high and the singing voice synthesizer unit 202 and the musical-sound synthesizer unit 203 are given. 207 is the input terminal of performance rate indicated value, and a sequencer makes a performance rate quickly or late with the value given here.

[0077] Since the thing of the same configuration as the gestalt 1 of operation is used for the singing voice synthesizer unit 202 with the gestalt of this operation, if there are transposition directions from an input terminal 206 and a sequencer 201 changes pitch indicated value, the address which reads the wave memory 105a and 105b will change.

[0078] However, since this does not change the quality of the data point to read, in spite of not /Carrying out, a singing voice composition machine is not the deteriorating thing to transpose.

[0079] Since the timing information given to the read-out machines 106a and 106b, the EMBE formation machines 107a and 107b, etc. only changes and quality of the data point read as well as the case of transposition is not changed even if performance rate indicated value changes similarly, in spite of not /Carrying out, a singing voice synthesizer unit is not the deteriorating thing to transpose.

[0080] As mentioned above, even if it can play music and operates contest a key etc. at the time of music playback of karaoke etc., without degrading the quality of singing voice even if it changes a performance rate and a musical interval by forming the performance data memory 200 in which performance information is stored, a sequencer 201, the singing voice synthesizer unit 202 by the gestalt 1 of operation, and the musical-sound synthesizer unit 203, musical quality does not deteriorate.

[0081]

[Effect of the Invention] ASAINA which outputs by turns the performance information which consists of a phoneme, a pitch, sound volume, and pronunciation timing to the 1st and 2nd output sequences as mentioned above according to this invention, the musical interval which 4s pitches and

pronunciation timing are directed by the 1st of ASAINA, and the 2nd output sequence, and outputs the 1st and 2nd musical interval change values to the 1st and 2nd outputs, respectively -- difference -- with a generator a pitch, a phoneme, and pronunciation timing are directed according to the 1st of ASAINA, and the 2nd output sequence, respectively -- having -- and a musical interval -- difference -- with the 1st [to which a musical interval change value is directed with the 1st and 2nd outputs of a generator respectively], and 2nd read-out machines The 1st and the 2nd EMBE formation machine which sound volume and pronunciation timing are directed, respectively and generate wave envelopment data according to the 1st of ASAINA, and the 2nd output sequence, The 1st and 2nd wave memory from which two or more phonemes and the voice wave of a pitch are beforehand memorized, and data are read with the 1st and 2nd read-out vessels, respectively, the 1st and 2nd level converters which carry out the level conversion of the output of the 1st and 2nd read-out machines with the output of the 1st and 2nd EMBE formation machines, respectively, and the adder adding the output of the 1st and 2nd level converters -- **** -- by things It reads according to an individual by turns for every phoneme, and the voice wave of the specified phoneme/pitch is outputted smoothly, reading and carrying out cross fade with a vessel. except for the cross fade section -- a fundamental tone PCM - - since it carries out asymptotic [of the pitch of a front phoneme] to a back pitch gradually at the time of cross fade when the high singing voice of quality can be compounded and there is a pitch difference by the front phoneme and the back phoneme further, since it remains as it is, change of a pitch becomes smooth, and natural singing voice can be compounded.

[0082] Moreover, the performance data memory which has memorized a musical instrument and the performance information on a song, The sequencer which outputs performance directions information while controlling timing according to the performance information, transposition input, and performance rate input which were read from performance data memory, the musical-sound synthesizer unit which compounds musical sound according to performance directions of a

sequencer, the singing voice synthesizer unit which compounds singing voice according to performance directions of a sequencer, and the adder adding the output of a musical-sound synthesizer unit, and the output of a singing voice synthesizer unit -- **** -- by things By being able to perform music playback with singing voice with high quality, also when changing the musical interval and rate to perform, and applying this to the music regenerative apparatus of synthesizer karaoke The synthesizer karaoke to which the quality of a back chorus does not fall even if a performance rate and a musical interval change can be offered, and it has the practically excellent effectiveness.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
 - 2.**** shows the word which can not be translated.
 - 3.In the drawings, any words are not translated.
-

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The block diagram showing the configuration of the singing voice synthesizer unit in the gestalt 1 of operation of this invention

[Drawing 2] The explanatory view showing the actuation and the configuration of a read-out machine in the gestalt of this operation

[Drawing 3] the musical interval in the gestalt of this operation -- difference -- the explanatory view explaining actuation of a generator

[Drawing 4] The explanatory view of the data storage condition of the wave memory in the gestalt of this operation

[Drawing 5] The explanatory view of loop-formation read-out of the wave memory in the gestalt of this operation

[Drawing 6] The explanatory view of cross fade processing of the singing voice composition in the gestalt of this operation

[Drawing 7] The explanatory view of ASAINA in the gestalt of this operation of operation

[Drawing 8] The block diagram showing the configuration of the music regenerative apparatus in the gestalt 2 of operation of this invention

[Description of Notations]

104 ASAINA
105a, 105b Wave memory
106a, 106b Read-out machine
107a, 107b EMBE formation machine
108a, 108b Level converter
109 Adder
111 Musical Interval -- Difference -- Generator
200 Performance Data Memory
201 Sequencer
202 Singing Voice Synthesizer Unit
203 Musical-Sound Synthesizer Unit
204 Adder

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any
damages caused by the use of this translation.

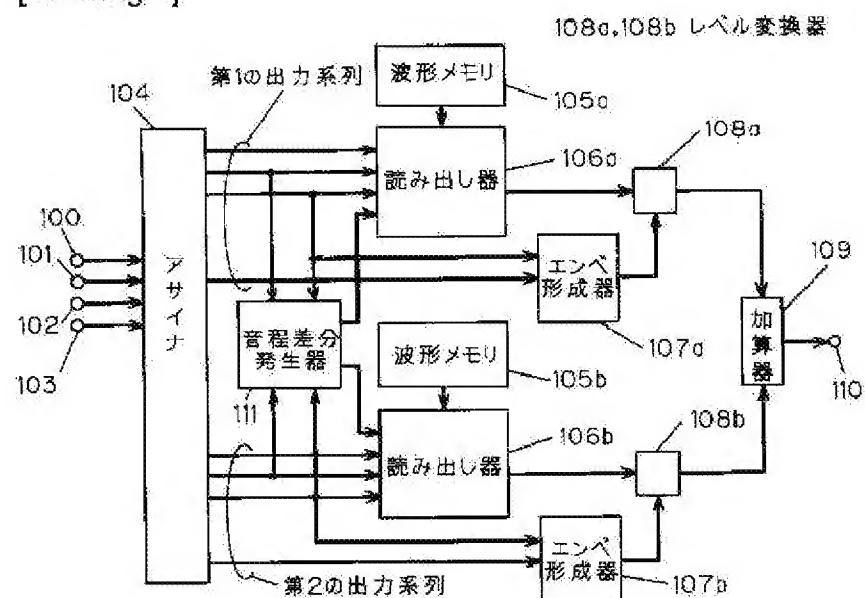
1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

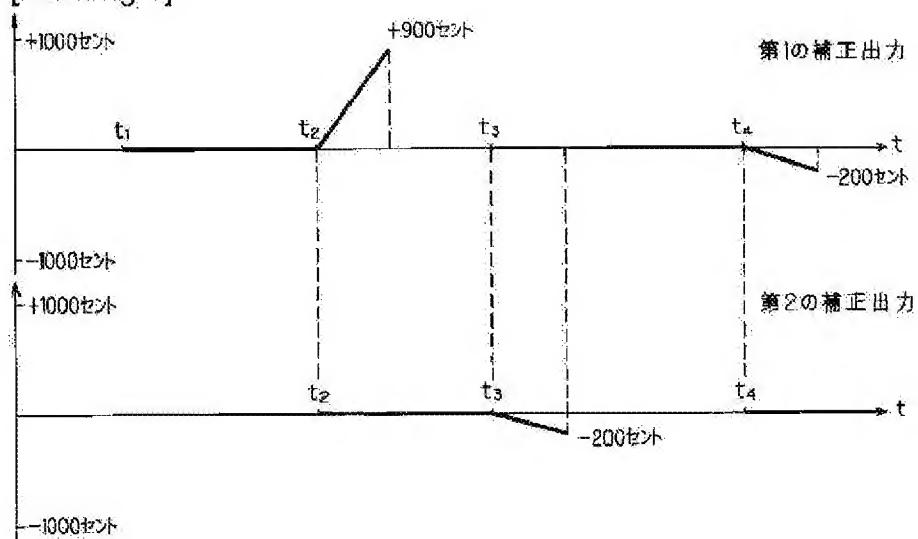
3. In the drawings, any words are not translated.

DRAWINGS

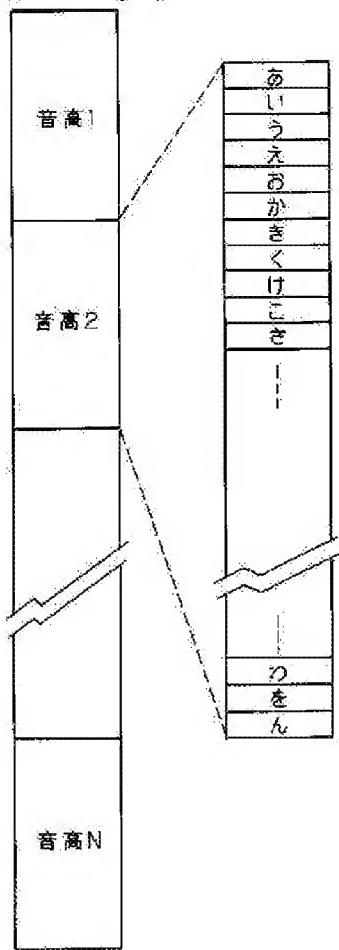
[Drawing 1]



[Drawing 3]



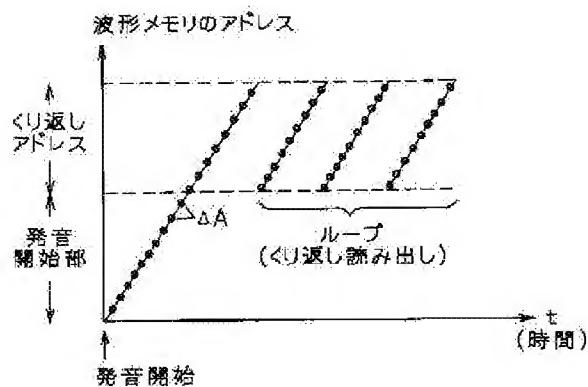
[Drawing 4]



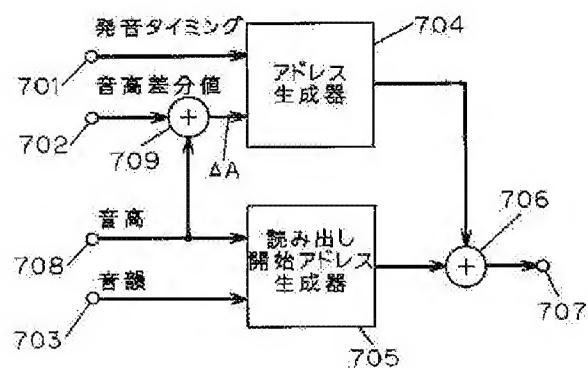
[Drawing 2]

706,709 加算器

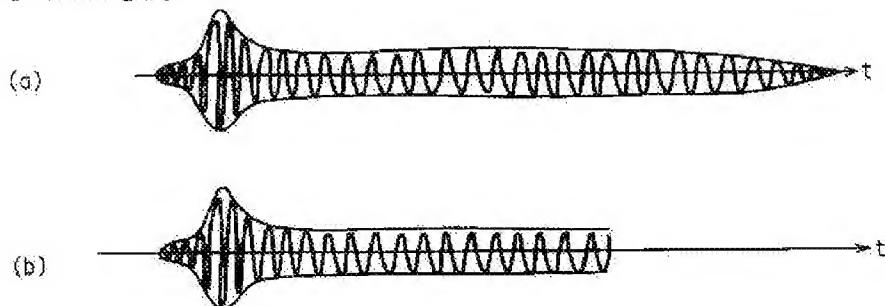
(a)



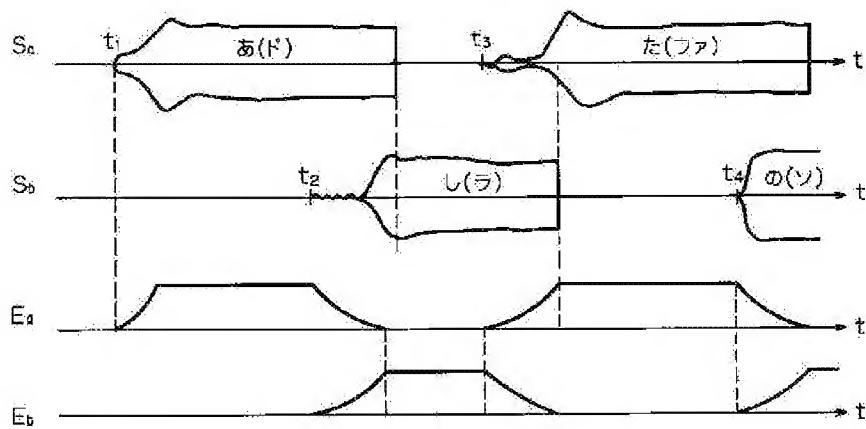
(b)



[Drawing 5]



[Drawing 6]



[Drawing 7]

アサイナ 入力	時間	t_1		t_2		t_3		t_4
	音韻	あ		し		た		の
音高	ド			ラ		ソ		ファ
発音	ON			ON		ON		ON
音量	100			100		100		100

第1の 出力系列	音韻	あ		た
	音高	ド		ソ
発音	ON		ON	
音量	100			100

第2の 出力系列	音韻	し		の
	音高	ラ		ファ
発音	ON		ON	
音量	100			100

[Drawing 8]

